Received: 22 December 2023 | *Revised: 4 February 2024* | *Accepted: 23 February 2024*
DOI: 10.37917/ijeee.21.2.25 *Early View | December 2025*

Open Access

Iraqi Journal for Electrical and Electronic Engineering
Original Article

Mobility Prediction Based on LSTM Multi-Layer Using GPS Phone Data

Nabaa Kareem Mhalhal*, Suhad Faisal Behadili
Department of Computer Science, College of Science, University of Baghdad, Baghdad, Iraq

Correspondance
*Nabaa Kareem Mhalhal
Department of Computer Science, College of Science, University of Baghdad,
Baghdad, Iraq
Email: naba.kareem2201m@sc.uobagdad.edu.iq

Abstract
Precise Prediction of activity location is an essential element in numerous mobility applications and is especially necessary for the development of tailored sustainable transportation systems. Next-location prediction, which involves predicting a user's future position based on their past movement patterns, has significant implications in various domains, including urban planning, geo-marketing, disease transmission, Performance wireless network, Recommender Systems, and many other areas. In recent years, various predictors have been suggested to tackle this issue, including state-of-the-art ones that utilize deep learning techniques. This study introduces a robust Model for predicting the future location path of a user based on their known previous locations. The study proposes the use of a Long Short-Term Memory (LSTM) prediction scheme, which is well-suited for learning from sequential data; then a fully connected neuron is employed to decrease the sparsity of the data, resulting in accurate predictions for the path of the user's next location. The suggested strategy demonstrates superior prediction accuracy compared to a state-of-the-art method, with improvements of up to a loss error of 0.002 based on real-life datasets (Geolife). The results demonstrate that the reliability of forecasts is excellent, indicating the accuracy of the predictions.

Keywords
Deep Learning, Next Location, Long Short-Term Memory (LSTM), Prediction.

I. Introduction

To accommodate the increasing prevalence of smart cities, autonomous vehicles, and the Internet of Things (IoT) in our daily lives, the future must be equipped with high-speed cellular networks that are both efficient and capable of delivering excellent quality of service (QoS) and quality of experience (QoE) [1]. Individual mobility refers to the act of relocating from one geographic area to another and is characterized by a collective of persons or mobile individuals [2–4]. The field of user mobility research has yet to be thoroughly investigated, and predicting individual user movement is a complicated and multifaceted task. Current individual mobility prediction models underperform with only 93% accuracy, creating a critical gap in the ability to reliably anticipate user movements. Addressing this gap requires a robust theoretical framework and practical development to enhance the prediction precision. Despite being essential, location prediction remains a formidable challenge that has yet to be resolved entirely. It is well acknowledged that individual movements on a collective level have a highly consistent pattern that may be characterized using statistical distributions [1, 5–8]. The primary objective of this study is to enhance services in various major domains through the prediction of next-location. These domains include:

1. It is essential in several applications, such as providing travel suggestions and optimizing routes, detecting potential public emergencies at an early stage, estimating urban emissions, offering location-based advertisements and geo-marketing, and making friend recommendations on social network platforms.

2. Estimating the future position of a mobile user utilizing its trajectory information is a crucial aspect in developing intelligent applications that can assist users without requiring manual input [5].

3. The estimation and prediction of user mobility in an ad hoc wireless network will have an advantageous effect on ensuring the quality of service (quality of service) and quality of experience (QoE) inside the network [6].

4. Also, the forthcoming 5G and B5G mobile communication technologies are expected to offer significant improvements in terms of reduced latency and delay [5].

5. The mobility of customers in their everyday activities is greatly influenced by the quality of telecom services. This is particularly important in vehicular ad hoc networks (VANETs), as it contributes to addressing issues such as traffic congestion, handovers, and activity routing protocol [5].

6. In the field of urban science, the use of mobility user's trajectory data helps researchers comprehend urban dynamics, pinpointing areas of high activity and enhancing the distribution of resources and public services [7].

Artificial Neural Networks (ANN) can be used as a reliable and efficient method to predict user mobility. This nontraditional approach utilizes users' past patterns and behaviors to forecast future locations with more accuracy and trustworthiness [8]. Nevertheless, the existing prediction models developed by various researchers have been shown to be more robust in delivering a reliable and comprehensive forecast for subsequent location determination [9].

The literature study section below provides a comprehensive analysis of the extensive research conducted on forecasting a user's future location. Most of the current research is primarily on the geographical component of a user movement, such as the journey route, while ignoring the temporal part, including factors like travel time, arrival time, and duration [5].

However, predicting the future location of a user is particularly difficult, especially from the perspective of a network [10]. The main contributions of this research are the efficient design of LSTM Models that predict the user's next location. The suggested method provides numerical data and visualization results that prove the performance of user mobility prediction.

The subsequent sections of the article are structured in the following manner: Section II provides an overview of the relevant literature. Section III offers an overview of fundamental concepts, terminologies, and outlines of the utilization of the LSTM neural network for the purpose of predicting the user's next location. Section IV conducts an assessment of the suggested prediction model. Finally, Section V provides the concluding remarks of the paper.

II. Related Works

The rapid growth of mobile devices has led millions of users to generate enormous quantities of trajectory data, including Global Positioning System (GPS) log data, [11]. Trajectory data encompass extensive details regarding travel habits and the overall traffic dynamics across a network. This data provides novel possibilities for uncovering and predicting both individual and collective mobility patterns of users within urban road networks [12].

Next-location prediction seeks to predict the specific location that an individual will visit next, utilizing their previously visited locations as a basis. Different methodologies have been employed in previous studies to anticipate the next location, including deep learning (DL) [13]. The future of deep learning is promising, as it excels in effectively handling large volumes of diverse data, which is particularly essential in the current era of intelligent sensors that gather substantial amounts of information [14].

Within this set of methods, DL approaches, such as recurrent neural network (RNN), Long Short-Term Memory (LSTM), and bidirectional LSTM (BiLSTM), are the most common choices in the literature. For example, Yi Bao et al [15] utilized the BiLSTM-CNN (Bidirectional Long Short-Term Memory-Convolutional Neural Network) model to predict the next locations using geotagged social media and got high accuracy that has significant relevance in area-based prediction, and recommendation reach to an 80% accuracy with latitude/longitude coordinates without take consideration time. Another study, such as Jie Sun and Jiwon Kim [16] used a combination of hybrid LSTM and sequential LSTM models. These models analyze patterns in the spatial and temporal sequences of trajectory data and their interconnections in order to predict both future locations and travel durations simultaneously using Bluetooth data. While Feng Li et al [17] presented an unusual method to optimize the spectrum allocation for the Internet of Vehicles (IoV) utilizing a combination LSTM with RNN to predict the next location positions of IoV nodes based on individual user movement.

Furthermore, Ye Hong et al [18] employed a multi-head self-attentional (MHSA) neural network to analyze previous location visits, visit time, activity duration, and nearby land use functions. The purpose was to predict an individual subsequent location based on learned location transition patterns.

In this study, we are going to build a robust model to deal with the problem of the next location for individuals. In the next section, we explain the fundamentals of the LSTM models used in deep learning with a brief explanation, and then we explain the designed model.

III. Research Methodology

This section provides an overview of the various components of the suggested model and presents a comprehensive explanation of the proposed deep learning model architecture.

A. The Fundamentals of The LSTM Model

LSTM, is a distinct type of (RNN) that excels at capturing and understanding long-term dependencies in data [19], where the LSTM cell contains a cell state (C), a Forget gate, an Input gate, and an output gate. Cell state: The Cell state handles the long-term memory, which allows storing the data related to previous cells in the LSTM cell. Forget gate, which is below the cell state, modifies it, whereas the input modulation gate adjusts the cell state; Fig. 1 illustrates the capabilities of LSTM.

where C is the Cell state, f is the forget gate, i is the input gate, o is the output gate, h is the hidden state, σ is the sigmoid function, tanh is the activation function [19]. So, LSTM cell can be explained as:

1. Forget gate is multiplied by the previous cell state and added to the input gate multiplied by the current cell state, which is calculated as in [19]:

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t \tag{1}$$

2. Forget gate forgets the output when multiplied by zero and stores the information in the cell state by one to a position in the matrix. In Eq 2, the sigmoid function pertains to the weighted function of the previous hidden state and input [19]:

$$f_t = \sigma\left(w_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{2}$$

3. Input gate: Activation functions are essential parts of every gate. This gate handles the type of information to enter the cell state. The summation of the previous cell state is the equation of the cell state, where the sigmoid function cannot forget the memory and can only add memory within the range of [0,1]. If a float number is added with the range [0,1], it can never be zero/forget. So, the tanh activation function is applied to the weighted input with the range [-1,1], which permits the cell state to forget the memory as in 3 and in 4 [19]:

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{3}$$

$$\hat{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{4}$$

4. Output gate: This gate handles the information about which value is to be moved from the matrix to the next hidden state from all the possible values, as in 5 and 6 [19]:

$$o_t = (\sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \tag{5}$$

$$h_t = o_t * \tanh\left(C_t\right) \tag{6}$$

B. The Proposed Model

Actually, in order to address the issue of the vanishing gradient problem in handling long-term sequential data, this work employs an LSTM recurrent neural network [20]. Typically, this strategy yields greater predicting accuracy compared to the conventional approach of training the Neural Network using the entire dataset (Geolife) [21]. Hence, several libraries were employed to mitigate the complexity of building the anticipated Neural Network. The most significant libraries for selecting the optimal configuration for the examined dataset are TensorFlow, Keras, Sklearn, NumPy, and Pandas [22]. Then, the LSTM Multi-Layer Models network in TensorFlow has been trained and tested. We can explain this as follows:

- Phase 1: To predict the next location, the latitude and longitude coordinates are extracted from each line of the file. Sequences of a specified length are generated from the list of latitudes and longitudes and saved. The subsequent coordinate in each sequence is kept as a label in the label's variable. Then The lists of sequences data and labels are converted to NumPy arrays for compatibility with many machine learning frameworks. The array X holds the sequences, while y holds the corresponding next coordinates.

- Phase 2: The output of phase 1 is the input to phase 2, a one hundred neuron LSTM neural network is employed for the first layer, followed by a fifteen neuron of LSTM neural network for the second layer.

- Phase 3: The output from the previous phase goes through a dense layer consists of two neuron since we predict latitudes and longitudes. Fig 2 explains the flow chart of the suggested model.

Additionally, in order to enhance the accuracy of the model various methodologies were employed on the DL models, including early stopping. The primary objective during model training is to minimize the loss function, namely the Mean Square Error (MSE). Early stopping is a technique used to terminate the training process when the monitored parameter refuses to show improvement. Subsequently, the early stopping technique is successful in consuming the required epochs training time, and to avoid overfitting. Meanwhile, the metric is assessed at the conclusion of every epoch. The application of this technique gives an MSE value of 0.002, which could be considered as the optimal result of the model [22].

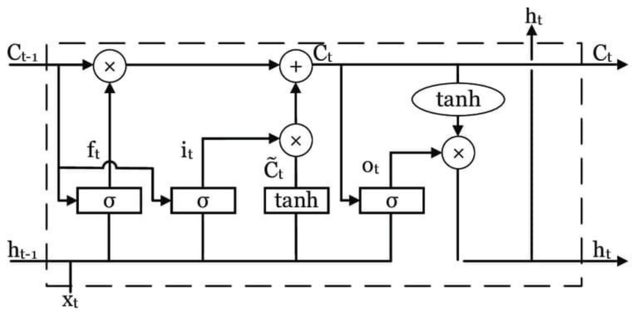


Fig. 1. LSTM neural network basic

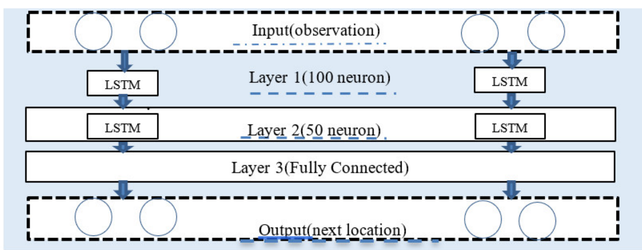
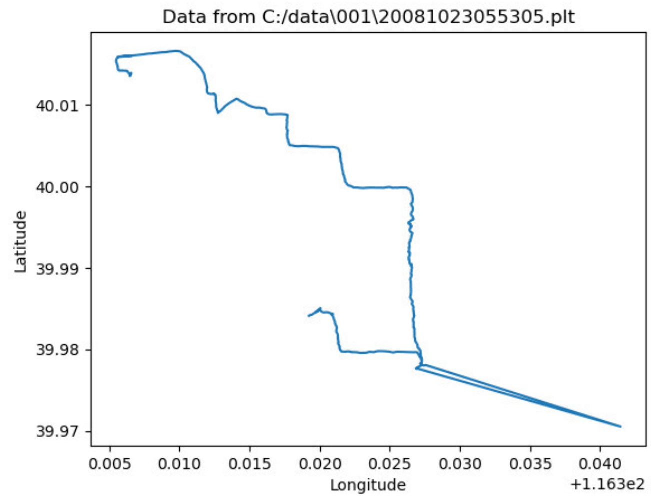


Fig. 2. Flow chart of the proposed model

	Time	latitude	Longitude	Altitude	user
0	2008-10-23 02:53:04	39.984702	116.318417	492.0	0
1	2008-10-23 02:53:10	39.984683	116.318450	492.0	0
2	2008-10-23 02:53:15	39.984686	116.318417	492.0	0
3	2008-10-23 02:53:20	39.984688	116.318385	492.0	0
4	2008-10-23 02:53:25	39.9847655	116.318263		
…	…	…	…	…	…
1853	2009-01-22 06:12:54	39.887149	116.641007	18.0	0

(a) Raw data set



(b) Trajectory from one user

Fig. 3. Detailed of the dataset

IV. Results and Dataset

A. Dataset The work has generally applied the DL model to the real-life (GeoLife) GPS Trajectories dataset. The GPS trajectory information was obtained from the Geolife study conducted by 182 users at Microsoft Research Asia [21]. The dataset captured a diverse range of individuals participating in various outdoor activities, including everyday tasks like commuting to work, as well as recreational and sports pursuits such as shopping, sightseeing, dining, hiking, and cycling. It is widely employed in diverse study fields, such as mobility pattern analysis, user activity recognition, location-based social networks, location privacy, and location recommendation. The examined dataset contains the GPS trajectory data, which consists of a series of time-stamped points, denoted as $Tr = \{tr_1, tr_2, \ldots, tr_{rn}\}$, obtained from an urban area or city. Each trajectory comprises a sequence of m recordings. The sequence is denoted as $t = t_1, t_2, \ldots, t_m$. Each record t is a tuple that comprises t = {sequence number, time as string (including year, month, day, hour), latitude, longitude, altitude, number of user} [21]. As shown in Fig. 3, the raw data set is displayed, and trajectories have been plotted.

A. *Experimental Set-up*

In this work, the purpose behind applying the LSTM Multi-Layer Models network is to predict the individual's next location paths. The deep learning models were implemented on a Lenovo system consisting of Intel(R) Core (TM), i5-8250U Gen processor @ 1.60GHz, 1.80 GHz, and 16 GB RAM. This system runs on Windows 10 Pro 64-bit operating system. The experiments are conducted using Keras library with Tensorflow library as the backend in Jupyter Notebook environment. We implemented the deep learning method that is the LSTM Multi-Layer on an actual life data set (GeoLife) [21] as:

1. Preprocessing dataset: The dataset contains 182 folders. Each folder in this collection contains the PLT-formatted GPS log files that belong to a certain user.

   Every PLT file has a unique trajectory with a starting time assigned to it. The raw data contains 6 columns of data. Once the raw data is extracted from the server, it is to be analyzed and the features that are required for the next location are to be selected. The first six lines in each folder must be drop. Are presumed to be headers, are skipped. Requires the manipulation of the acquired data and its representation through graphical visualization [21]. Given that each line in the data list consists of comma-separated values, with the first value denoting latitude and the second value denoting longitude, these lines extract and convert these values into floating-point data types, resulting in two lists: latitudes and longitudes as in the Algorithm 1:

2. Model training: Each of the 15 participants has GPS data records over a maximum period of 65 days (approx. 9 weeks), although some participants had less than 65 days. On average, participants' GPS data were available on 60 days with a minimum of 50 days. We train the proposed model utilizing 80% of data for each selected individual while reserving the remaining 20% for the testing set.

3. Hyperparameters: Mobility models that forecast the user future trajectory or path (also known as regression tasks) [8]. So to optimize the proposed model, the loss function Mean Square Error (MSE) has been chosen. Since, it makes it possible to deal with coordinates, especially, we don't need to examine a negative input. Also, a suitable activation function is the Rectified Linear Unit (ReLU) is the identity function for positive input and zero for negative input [23]. While the optimization technique that has been used is the Adaptive Moment Estimation (Adam), which is an adaptive optimization algorithm that is very effective and is commonly used to replace the conventional stochastic gradient reduction technique [24]. Hence, Table I shows the Hyperparameters of the proposed model.

Algorithm 1: The Pre-processing dataset.

```
Input  : Dataset
Output : Obtain two arrays (X,Y)
Initialize data path
for loop to open the file for reading do
    for loop to Skip the first six lines (assumed headers) do
        Read the remaining lines, extracting latitude
          and longitude coordinates from each line.
    end
end
Normalization arrays(X, Y)
return two arrays(X, Y)
```

B. Evaluation Metrics

As the final phase, we evaluated the prediction model using an evaluation score for the loss function with the MSE functions as a statistical measure that quantifies the average difference between the predicted values and the actual values in a dataset. The quantity being referred to is the standard deviation of the projected errors. The concentration of information on the best fit line is determined by this factor, which is defined in 7 [25]:

TABLE I.
HYPERPARAMETERS OF PROPOSED MODEL

Epochs	30
Batch size	64
Optimization algorithm	Adam
Loss function	MSE
Number of layers	3
Number of neurons	152
Activation function	ReLU

$$MSE = \frac{1}{m}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{7}$$

Root Mean Square Error (RMSE): It is the standard deviation of the predicted errors as in 8 [25].

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{8}$$

The term Mean Absolute Error (MAE) refers to the average difference between the actual values and the predicted values, calculated by taking the absolute value of each difference and then finding the average. It denotes the mean significance of the errors and is defined in 9 [25]:

$$MAE = \frac{1}{M}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{9}$$

Mean Absolute Percentage Error (MAPE) is a metric used to measure the accuracy of a predictive by calculating the average percentage difference between the forecasted values and the actual values. The calculation involves dividing the total sum of absolute mistakes by the number of individuals, resulting in the average of those individual absolute errors. The suggested models forecast deviation from its corresponding output is determined and defined in 10 [25]:

$$MAPE = \frac{1}{M}\sum_{i=1}^{n}\frac{y_i - \hat{y}_i}{y_i} \tag{10}$$

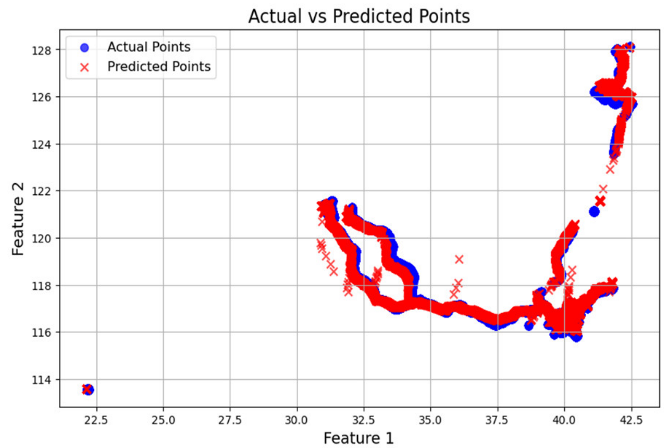
Where y_i is the actual value, $\hat{y}_i$ is the predicted value, and m is the total number of observations. MSE, MAE, and MAPE are suitable measurements for computing prediction sequential model accuracy and are the main criteria to define if the research is regarding the prediction model [25]. All those measurements have the best range, which is (0).

C. Results

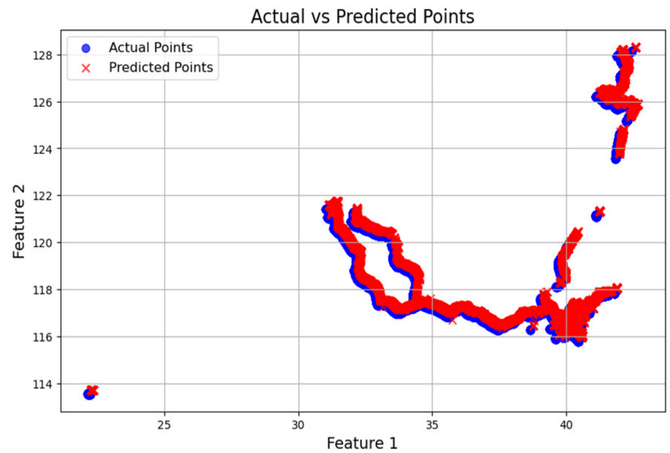
1) Comparison with Two Traditional Deep Learning Approaches and with Other Similar Work of the Same Attitudes

In Table 2, the performance of the proposed model was compared with that of traditional approaches for the same Hyperparameters. We use the data from 15 participants for training and testing. For each participant, 80% is used as training data, and the remaining 20% is used as testing data. The methods used for comparisons are as follows:

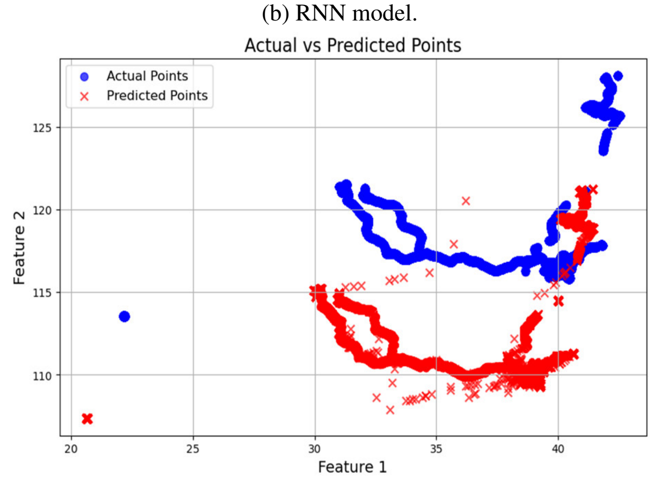
- Recurrent Neural Networks (RNNs) refer to an extension of the standard feed-forward neural network that is capable of processing input sequences of varying lengths [19].
- Gated Recurrent Units (GRU) refer to a version of LSTM, consists of gate functions to address the issues of gradient vanishing and gradient exploding in RNN. For more detailed about these models, see [19].
- The performance of the suggested model was evaluated and compared with work of the researcher Ida Nurhaida et.al [26].



(a) Proposed model.



(b) RNN model.



(c) GRU model.

Fig. 4. Visualization results where the predicted trajectories (red) actual trajectories (blue)

The results clearly indicate that proposed model outperforms the state-of-the-art and competitive two traditional deep learning approaches in terms of MSE, MAE and MAPE. Fig. 4 compared visualization results of the proposed model with other models, Fig. 5 shows the Visualization results of the proposed model for evaluation metrics.

V. Conclusion

The study effectively utilized the GeoLife GPS Trajectories dataset to predict the individual's mobility. The next location path has been excluded using the LSTM Multi-layer approach prediction model on this dataset. The evaluation of prediction using the LSTM method yielded a score of 30 for the epoch parameter, and 64 for the batch size parameter, indicating a high level of accuracy. Additionally, the prediction error is measured by the mean square error (MSE), achieving a score of 0.002. Finally, according to the concluded results, potential future research endeavors may involve investigating the feasibility of integrating other data sources, real-time traffic information, or weather patterns into the LSTM model. This has the potential to improve the precision and level of detail in future location predictions.

Conflict of Interest

The authors have no conflict of relevant interest to this article.

References

[1] X. Wang, M. Fang, Z. Zeng, and T. Cheng, "Where would i go next? large language models as human mobility predictors," *arXiv:2308.15197*, 2023.

[2] M. K. AlQaysi, S. F. Behadili, and A. Salam, "Spatiotemporal modeling in wireless communication networks," *Baghdad Science Journal*, vol. 20, no. 3, pp. 0904–0904, 2023.

[3] M. K. Alqaysi and S. F. Behadili, "Gravity model for flow migration within wireless communication networks," *Iraqi Journal of Science*, vol. 63, no. 10, pp. 4474–4487, 2022.

[4] Y. Hong, H. Martin, and M. Raubal, "How do you go where? improving next location prediction by learning travel mode information using transformers," in *Proceedings of the 30th International Conference on Advances in Geographic Information Systems*, (Seattle Washington), pp. 1–10, 2022.

[5] B. Mo, Z. Zhao, H. N. Koutsopoulos, and J. Zhao, "Individual mobility prediction in mass transit systems using smart card data: An interpretable activity-based hidden markov approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 12014–12026, 2021.

[6] H. R. Pamuluri, *Predicting User Mobility using Deep Learning Methods*. Ms.c thesis, Blekinge Institute of Technology, Karlskrona, Sweden, 2020.

[7] D. do Couto Teixeira, J. M. Almeida, and A. C. Viana, "On estimating the predictability of human mobility: the role of routine," *EPJ Data Science*, vol. 10, no. 1, p. 49, 2021.

[8] N. Bahra and S. Pierre, "A hybrid user mobility prediction approach for handover management in mobile networks," *Telecom*, vol. 2, no. 2, pp. 199–212, 2021.

[9] M. Luca, G. Barlacchi, B. Lepri, and L. Pappalardo, "A survey on deep learning for human mobility," *ACM Computing Surveys (CSUR)*, vol. 55, no. 1, pp. 1–44, 2021.

[10] M. Luca, L. Pappalardo, B. Lepri, and G. Barlacchi, "Trajectory test-train overlap in next-location prediction datasets," *Machine Learning*, vol. 112, no. 11, pp. 4597–4634, 2023.

[11] B. Mo, Z. Zhao, H. N. Koutsopoulos, and J. Zhao, "Individual mobility prediction: an interpretable activity-based hidden markov approach," *arXiv:2101.03996*, 2021.

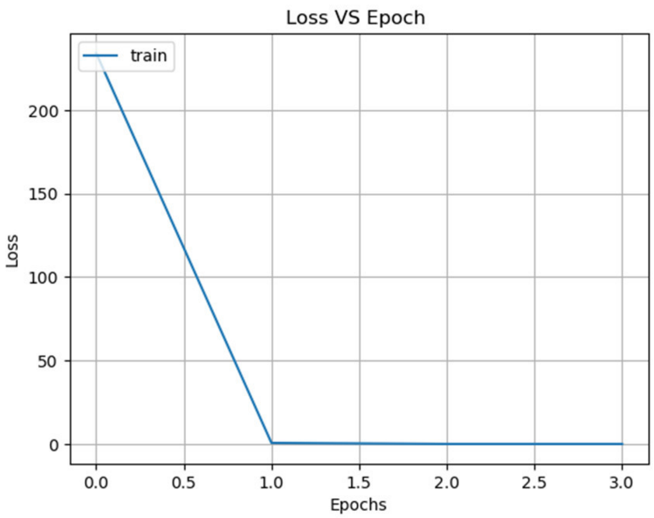
[12] P. Chen, W. Shi, X. Zhou, Z. Liu, and X. Fu, "Stlp-gsm: a method to predict future locations of individuals based on geotagged social media data," *International Journal of Geographical Information Science*, vol. 33, no. 12, pp. 2337–2362, 2019.

[13] M. Yan, S. Li, C. A. Chan, Y. Shen, and Y. Yu, "Mobility prediction using a weighted markov model based on mobile user classification," *Sensors*, vol. 21, no. 5, p. 1740, 2021.

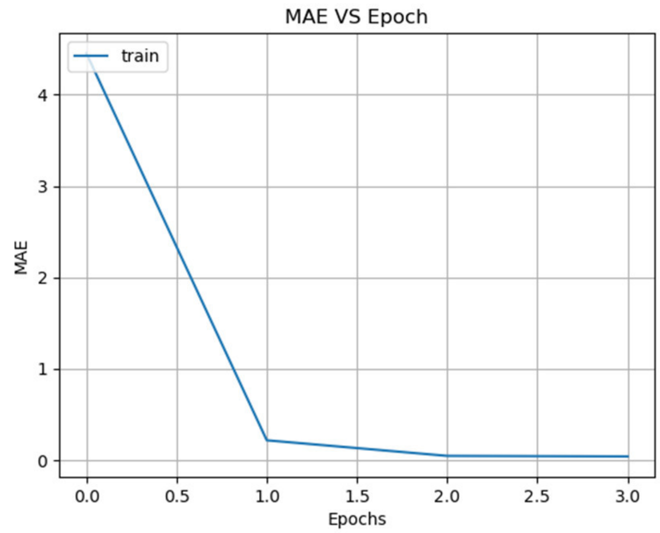
[14] H. N. Bhandari, B. Rimal, N. R. Pokhrel, R. Rimal, and K. R. Dahal, "Lstm-sdm: An integrated framework of lstm implementation for sequential data modeling," *Software Impacts*, vol. 14, p. 100396, 2022.

[15] Y. Bao, Z. Huang, L. Li, Y. Wang, and Y. Liu, "A bilstm-cnn model for predicting users' next locations based on geotagged social media," *International Journal of Geographical Information Science*, vol. 35, no. 4, pp. 639–660, 2021.

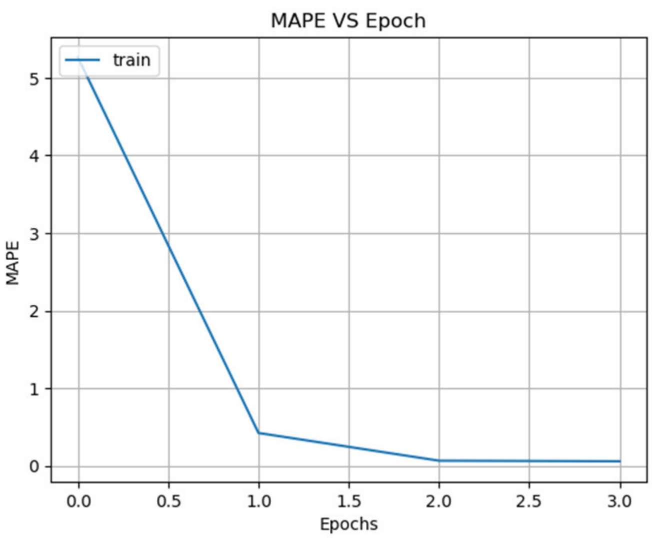
[16] J. Sun and J. Kim, "Joint prediction of next location and travel time from urban vehicle trajectories using



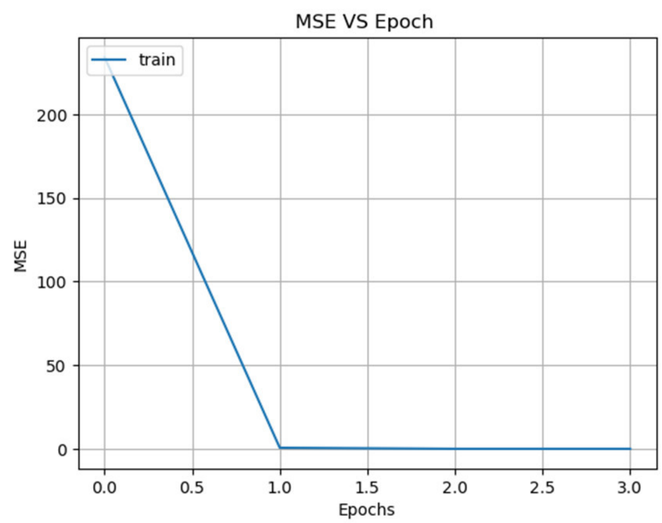
(a) Loss vs. epochs.



(b) MAE vs. epochs.



(c) MAPE vs. epochs.



(d) MSE vs. epochs.

Fig. 5. Visualization results of the proposed model for evaluation metrics.

long short-term memory neural networks," *Transportation Research Part C: Emerging Technologies*, vol. 128, p. 103114, 2021.

[17] F. Li, Z. Sun, K.-Y. Lam, L. Sun, B. Shen, and B. Peng, "Dynamic spectrum optimization for internet-of-vehicles with deep-learning-based mobility prediction," *Wireless Personal Communications*, vol. 137, p. 649–669, 2024.

[18] Y. Hong, Y. Zhang, K. Schindler, and M. Raubal, "Context-aware multi-head self-attentional neural network model for next location prediction," *Transportation Research Part C: Emerging Technologies*, vol. 156, p. 104315, 2023.

[19] F. M. Shiri, T. Perumal, N. Mustapha, and R. Mohamed, "A comprehensive overview and comparative analysis on deep learning models: Cnn, rnn, lstm, gru," *arXiv:2305.17473*, 2023.

[20] D. Yang, B. Fankhauser, P. Rosso, and P. Cudre-Mauroux, "Location prediction over sparse user mobility traces using rnns," in *Proceedings of the twenty-ninth international joint conference on artificial intelligence*, pp. 2184–2190, 2020.

[21] Y. Zheng, L. Wang, R. Zhang, X. Xie, and W.-Y. Ma, "Geolife: Managing and understanding your past life over maps," in *The Ninth International Conference on Mobile Data Management (mdm 2008)*, (Beijing, China), pp. 211–212, IEEE, 2008.

[22] K. B. Prakash, R. Kannan, S. A. Alexander, and G. Kanagachidambaresan, *Advanced Deep Learning for Engineers and Scientists*. Springer, 2021.

[23] S. R. Dubey, S. K. Singh, and B. B. Chaudhuri, "Activation functions in deep learning: A comprehensive survey and benchmark," *Neurocomputing*, vol. 503, pp. 92–108, 2022.

[24] R. Abdulkadirov, P. Lyakhov, and N. Nagornov, "Survey of optimization algorithms in modern neural networks," *Mathematics*, vol. 11, no. 11, p. 2466, 2023.

[25] U. Bhimavarapu, G. Battineni, and N. Chintalapudi, "Improved optimization algorithm in lstm to predict crop yield," *Computers*, vol. 12, no. 1, p. 10, 2023.

[26] I. Nurhaida, H. Noprisson, V. Ayumi, H. Wei, E. D. Putra, M. Utami, and H. Setiawan, "Implementation of deep learning predictor (lstm) algorithm for human mobility prediction," *International Association of Online Engineering*, vol. 14, no. 18, p. 132–144, 2020.